

# A NOVEL APPROACH FOR GLOBAL REGISTRATION OF MEDICAL IMAGES BASED ON LEARNING THE PRIOR APPEARANCE MODEL

Ayman El-Baz<sup>1</sup> and Georgy Gimel'farb<sup>2</sup>

<sup>1</sup>Bioimaging Laboratory, Bioengineering Department, University of Louisville, Louisville, KY, USA.

<sup>2</sup>Computer Science Department, University of Auckland, New Zealand.

## ABSTRACT

A new approach to align an image of a medical object with a given prototype (reference object) is proposed. Visual appearance of the images, after equalizing their signals, is modeled with a new variant rotation and scaling Markov-Gibbs random field with pairwise interaction model. Similarity to the prototype (reference object) is measured by a Gibbs energy of signal co-occurrences in a characteristic subset of pixel pairs derived automatically from the prototype (reference object) using our previous Linear Combination of Discrete Gaussians (LCDG) probabilistic model. An object is aligned by an affine transformation maximizing the similarity by using an automatic initialization followed by gradient search. Experiments confirm that our approach aligns complex objects better than popular conventional algorithms.

**Index Terms**— Markov-Gibbs Random Field (MGRF), Lung, Computed Tomography (CT).

## 1. INTRODUCTION

Image registration aligns two or more images of similar objects taken at different times, from different viewpoints, and/or by different sensors. The images are geometrically transformed to ensure their close similarity. Registration is a crucial step in many applied image analysis tasks, e.g. to fuse various data sources (such as computer tomography (CT) and MRI data in medical imaging) for image fusion, change detection, or multichannel image restoration; form and classify multi-band images in remote sensing; update maps in cartography, perform automatic quality control in industrial vision, and so forth. Co-registered medical images provide more complete information about the patient, help to monitor tumor growth and verify treatment, and allow for comparing the patient's data to anatomical atlases (For more detail about medical image registration see [1–3]).

Most of the known registration methods fall into two main categories: feature-based and area-based techniques [4]. Feature based techniques use sparse geometric features such as points, curves, and/or surface patches, and their correspondences to compute an optimal transformation. Area-based methods such as the classical least square correlation match directly image signals to avoid feature extraction [5].

More powerful mutual information (MI) based image registration [6, 7] exploits a probabilistic similarity measure that allows for more general types of signal deviations than correlation. The statistical dependency between two data sets is measured by comparing a joint empirical distribution of the corresponding signals in the two images to the joint distribution of the independent signals. Because the MI-based registration performs the best with multi-modal images [7], it is used in many of medical imaging applications. The joint distribution is estimated using Parzen windows [8, 9] or discrete histograms [10]. The main advantage of the MI is insensitivity to monotone variations of correspondence between the object and prototype signals, but the objects should be of almost identical shape apart from their affine geometrical and monotone signal transformations. The MI allows also

for some non-monotone signal correspondence variations although they may change the visual appearance too much and hinder registration accuracy (e.g., see [11] for more details about MI based image registration).

We consider a more general case of registering a medical object to a prototype with similar but not necessarily identical visual appearance under their relative 2D affine transformations and monotone variations of signal correspondences. The variations are suppressed by equalizing signals in the images. The co-registered equalized images are described with a characteristic subset of signal co-occurrence statistics. The description implicitly “homogenizes” the images, i.e. considers them as spatially homogeneous patterns with the same statistics. In contrast to the feature-based registration, the statistics characterize the whole object. In contrast to the conventional area-based techniques, similarities between the statistics rather than pixel-to-pixel correspondences are measured. Section 2 represents the equalized object and prototype images as samples of a generic Markov-Gibbs random field (MGRF) with pairwise pixel interaction. Gibbs potentials are analytically estimated from co-occurrence statistics for the prototype. Similarity between an affinely transformed object and the prototype is measured with a total Gibbs energy for a characteristic pixel neighborhood. A new algorithm for selecting the neighborhood for the MGRF model is introduced. After an automatic initialization, the affine transformation aligning the object with the prototype is found by the gradient search for the maximum Gibbs energy of the transformed object. Experiments in Section 3 confirm that our method is more efficient for complex medical objects than more conventional SIFT and MI based registration techniques.

## 2. MGRF BASED IMAGE REGISTRATION

### 2.1. Basic notation.

We denote  $\mathcal{Q} = \{0, \dots, Q-1\}$ ;  $\mathbf{R} = [(x, y) : x = 0, \dots, X-1; y = 0, \dots, Y-1]$ , and  $\mathbf{R}_p \subset \mathbf{R}$  a finite set of scalar image signals (e.g. gray levels), a rectangular arithmetic lattice supporting digital images  $g : \mathbf{R} \rightarrow \mathcal{Q}$ , and its arbitrary-shaped part occupied by the prototype, respectively. A finite set  $\mathcal{N} = \{(\xi_1, \eta_1), \dots, (\xi_n, \eta_n)\}$  of  $(x, y)$ -coordinate offsets defines neighbors  $\{((x+\xi, y+\eta), (x-\xi, y-\eta)) : (\xi, \eta) \in \mathcal{N}\} \wedge \mathbf{R}_p$  interacting with each pixel  $(x, y) \in \mathbf{R}_p$ . The set  $\mathcal{N}$  yields a neighborhood graph on  $\mathbf{R}_p$  to specify translation invariant pairwise interactions with  $n$  families  $\mathcal{C}_{\xi, \eta}$  of cliques  $c_{\xi, \eta}(x, y) = ((x, y), (x+\xi, y+\eta))$ . Interaction strengths are given by a vector  $\mathbf{V}^T = [\mathbf{V}_{\xi, \eta}^T : (\xi, \eta) \in \mathcal{N}]$  of potentials  $\mathbf{V}_{\xi, \eta}^T = [V_{\xi, \eta}(q, q') : (q, q') \in \mathcal{Q}^2]$  depending on signal co-occurrences; here  $T$  indicates transposition.

### 2.2. Image normalization.

To account for monotone (order-preserving) changes of signals (e.g. due to different illumination or sensor characteristics), the prototype

and object images are equalized using the cumulative empirical probability distributions of their signals on  $\mathbf{R}_p$ .

### 2.3. MGRF based appearance model.

In line with a generic MGRF with multiple pairwise interaction [12], the Gibbs probability  $P(g) \propto \exp(E(g))$  of an object  $g$  aligned with the prototype  $g^\circ$  on  $\mathbf{R}_p$  is specified with the Gibbs energy

$$E(g) = |\mathbf{R}_p| \mathbf{V}^\top \mathbf{F}(g)$$

where  $\mathbf{F}^\top(g)$  is the vector of scaled empirical probability distributions of signal co-occurrences over each clique family:  $\mathbf{F}^\top(g) = [\rho_{\xi,\eta} \mathbf{F}_{\xi,\eta}^\top(g) : (\xi, \eta) \in \mathcal{N}]$  where  $\rho_{\xi,\eta} = \frac{|\mathcal{C}_{\xi,\eta}|}{|\mathbf{R}_p|}$  is the relative size of the family and  $\mathbf{F}_{\xi,\eta}(g) = [f_{\xi,\eta}(q, q'|g) : (q, q') \in \mathcal{Q}^2]^\top$ ; here,  $f_{\xi,\eta}(q, q'|g) = \frac{|\mathcal{C}_{\xi,\eta;q,q'}(g)|}{|\mathcal{C}_{\xi,\eta}|}$  are empirical probabilities of signal co-occurrences, and  $\mathcal{C}_{\xi,\eta;q,q'}(g) \subseteq \mathcal{C}_{\xi,\eta}$  is a subfamily of the cliques  $c_{\xi,\eta}(x, y)$  supporting the co-occurrence ( $g_{x,y} = q, g_{x+\xi, y+\eta} = q'$ ) in  $g$ . The co-occurrence distributions and the Gibbs energy for the object are determined over  $\mathbf{R}_p$ , i.e. within the prototype boundary after an object is affinely aligned with the prototype. To account for the affine transformation, the initial image is resampled to the back-projected  $\mathbf{R}_p$  by interpolation.

The appearance model consists of the neighborhood  $\mathcal{N}$  and the potential  $\mathbf{V}$  to be learned from the prototype.

### 2.4. Learning the potentials.

The MLE of  $\mathbf{V}$  is proportional in the first approximation to the scaled centered empirical co-occurrence distributions for the prototype [12]:

$$\mathbf{V}_{\xi,\eta} = \lambda \rho_{\xi,\eta} \left( \mathbf{F}_{\xi,\eta}(g^\circ) - \frac{1}{Q^2} \mathbf{U} \right); (\xi, \eta) \in \mathcal{N}$$

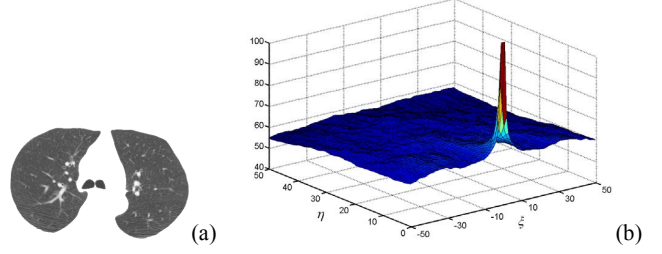
where  $\mathbf{U}$  is the vector with unit components. The common scaling factor  $\lambda$  is also computed analytically; it is approximately equal to  $Q^2$  if  $Q \gg 1$  and  $\rho_{\xi,\eta} \approx 1$  for all  $(\xi, \eta) \in \mathcal{N}$ . In our case it can be set to  $\lambda = 1$  because the registration uses only relative potential values and energies.

### 2.5. Learning the characteristic neighbors.

To find the characteristic neighborhood set  $\mathcal{N}$ , the relative Gibbs energies  $E_{\xi,\eta}(g^\circ) = \rho_{\xi,\eta} \mathbf{V}_{\xi,\eta}^\top \mathbf{F}_{\xi,\eta}(g^\circ)$  for the clique families, i.e. the scaled variances of the corresponding empirical co-occurrence distributions, are compared for a large number of possible candidates. Figure 1 shows a lung prototype and its Gibbs energies  $E_{\xi,\eta}(g^\circ)$  for 5000 clique families with the inter-pixel offsets  $|\xi| \leq 50; 0 \leq \eta \leq 50$ .

To automatically select the characteristic neighbors, let us consider an empirical probability distribution of the energies as a mixture of a large “non-characteristic” low-energy component and a considerably smaller characteristic high-energy component:  $P(E) = \pi P_o(E) + (1 - \pi) P_{hi}(E)$ . Because both the components  $P_o(E)$ ,  $P_{hi}(E)$  can be of arbitrary shapes, we closely approximate them with linear combinations of positive and negative Gaussians. For both the approximation and the estimation of  $\pi$ , we use the efficient EM-based algorithms introduced in [12] (subsequent steps of the approximation are shown in Fig. 2).

The intersection of the approximate mixture components gives an energy threshold  $\theta$  for selecting the characteristic neighbors as shown in Fig. 2(g):  $\mathcal{N} = \{(\xi, \eta) : E_{\xi,\eta}(g^\circ) \geq \theta\}$  where  $P_{hi}(\theta) \geq P_o(\theta)\pi/(1 - \pi)$ . The above example results in the threshold  $\theta = 28$  producing 173 characteristic neighbors shown in Fig. 3(a),(b)



**Fig. 1.** Lung image (a) and relative interaction energies (b) for the clique families in function of the offsets  $(\eta, \xi)$ .

together with the corresponding relative pixel-wise energies  $e_{x,y}(g^\circ)$  over the prototype:

$$e_{x,y}(g^\circ) = \sum_{(\xi,\eta) \in \mathcal{N}} V_{\xi,\eta}(g_{x,y}^\circ, g_{x+\xi, y+\eta}^\circ)$$

### 2.6. Appearance-based registration.

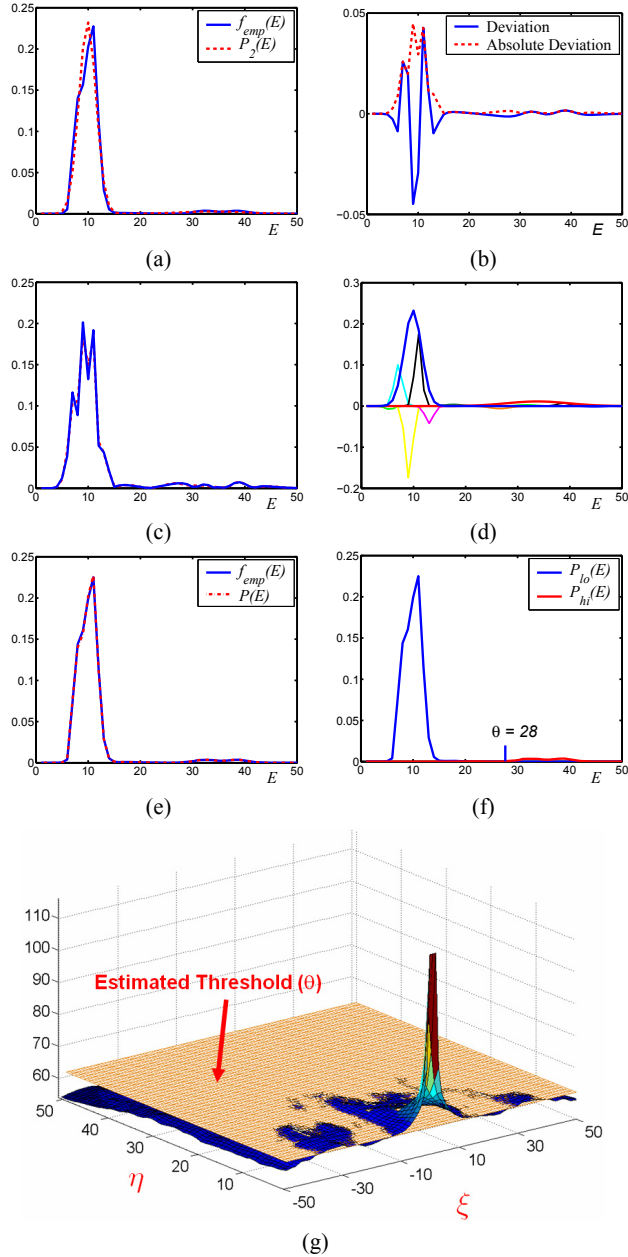
The object  $g$  is affinely transformed to (locally) maximize its relative energy  $E(g_a) = \mathbf{V}^\top \mathbf{F}(g_a)$  under the learned appearance model  $[\mathcal{N}, \mathbf{V}]$ . Here,  $g_a$  is the part of the object image reduced to  $\mathbf{R}_p$  by the affine transformation  $\mathbf{a} = [a_{11}, \dots, a_{23}]$ :  $x' = a_{11}x + a_{12}y + a_{13}$ ;  $y' = a_{21}x + a_{22}y + a_{23}$ . The initial transformation is a pure translation with  $a_{11} = a_{22} = 1$ ;  $a_{12} = a_{21} = 0$ , ensuring the most “energetic” overlap between the object and prototype. The energy for different translations  $(a_{13}, a_{23})$  of the object relative to the prototype is shown in Fig. 3(c); the chosen initial position  $(a_{13}^*, a_{23}^*)$  in Fig. 4(a) maximizes this energy. Then the gradient search for the local energy maximum closest to the initialization selects the six parameters  $\mathbf{a}$ ; Fig. 4(b) shows the final transformation aligning the prototype contour to the object.

## 3. EXPERIMENTAL RESULTS

Due to space limitations, we focus on low dose computed tomography (LDCT) of human lung commonly perceived as difficult for both the area- and feature-based registration. But the like results are obtained for several other types of complex objects (e.g., kidney images, zebra photos, starfish photos or brain images). We compare our approach to three popular conventional techniques, namely, to the area-based registration using MI [7] or normalized MI [10] and to the feature-based registration establishing correspondences between the images with SIFT [13]. Note that for both MI and NMI we used the implementations in ITK (Ver. 2.0). Results are shown in Fig. 4(b)–(e).

To clarify why the MI- or NMI-based alignment is less accurate, Fig. 5 compares the MI / NMI and Gibbs energy values for the affine parameters that appear at successive steps of the gradient search for the maximum energy. Both the MI and NMI have many local maxima that potentially hinder the search, whereas the energy is practically unimodal in these experiments. The SIFT-based alignment fails because it cannot establish accurate correspondences between similar lung and chest areas.

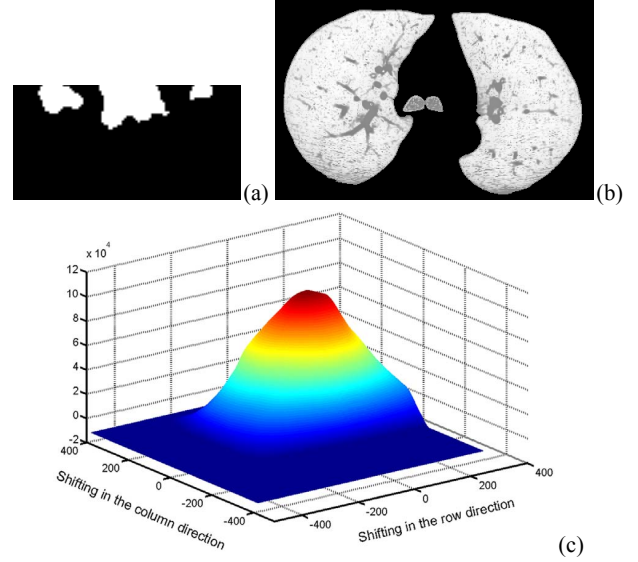
In the above example the object aligned with the prototype has mainly different orientation and scale. Figure 6 shows more diverse lungs and their Markov-Gibbs appearance-based and MI-based alignment with the prototype in Fig. 1(a). Visually, the back-projection of the prototype contour onto the objects suggests the better performance of our approach. To quantitatively evaluate the accuracy, masks of the co-aligned objects obtained by manual segmentation are averaged in Fig. 7. The common matching area is notably larger for our approach (96.8%) than for the MI-based registration (54.2%).



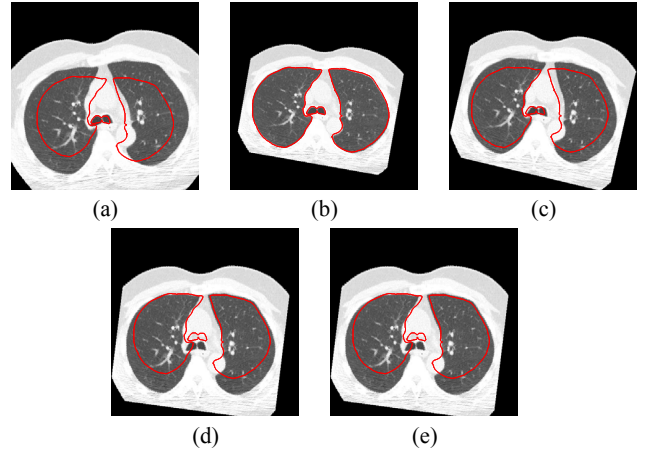
**Fig. 2.** Empirical density of the Gibbs energy  $f_{emp}(E)$  and the two estimated low-energy and high-energy dominant Gaussians  $P_2(E)$  (a); the alternate and absolute deviations between the empirical density and the two dominant components (b); the estimated mixture model of the absolute deviation (c); the final Gaussian components (d) in the joint estimated model (e) and the resulting low- and high-energy components (f) giving the energy threshold  $\theta$  to select the neighborhood system as shown in (g).

#### 4. CONCLUSIONS

In this paper we introduced a new approach to align an image of a medical object with a given prototype whose appearance is modeled with a new Markov-Gibbs random field with pairwise interaction model. Experimental results confirm that image registration based



**Fig. 3.** (a) Most characteristic 173 neighbors among the 5000 candidates (a; in white), (b) the pixel-wise Gibbs energies for the prototype under the estimated neighborhood, and (c) Gibbs energies for translations of the object with respect to the prototype.



**Fig. 4.** Initialization (a), our (b), MI-based (c), NMI-based (d), and SIFT-based (e) registration.

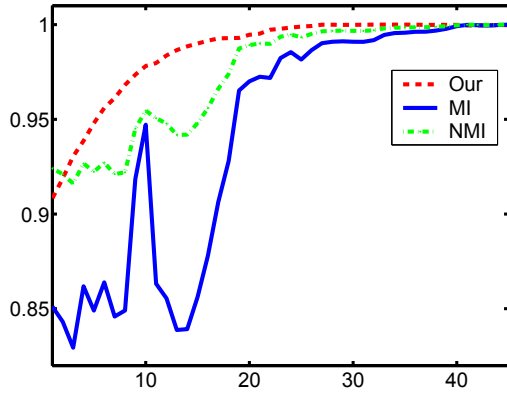
on our Markov-Gibbs appearance model is more robust and accurate than popular conventional algorithms. Due to the reduced variations between the co-aligned objects, our approach results in more accurate average shape models that are useful, e.g. in image segmentation based on shape priors.

As we mentioned in the experimental result section, the proposed approach is not only limited to lung images but also is suitable for registering zebra photos, starfish photos, and brain images.

**Acknowledgement:** This research work has been supported by Wallace H. Coulter Foundation.

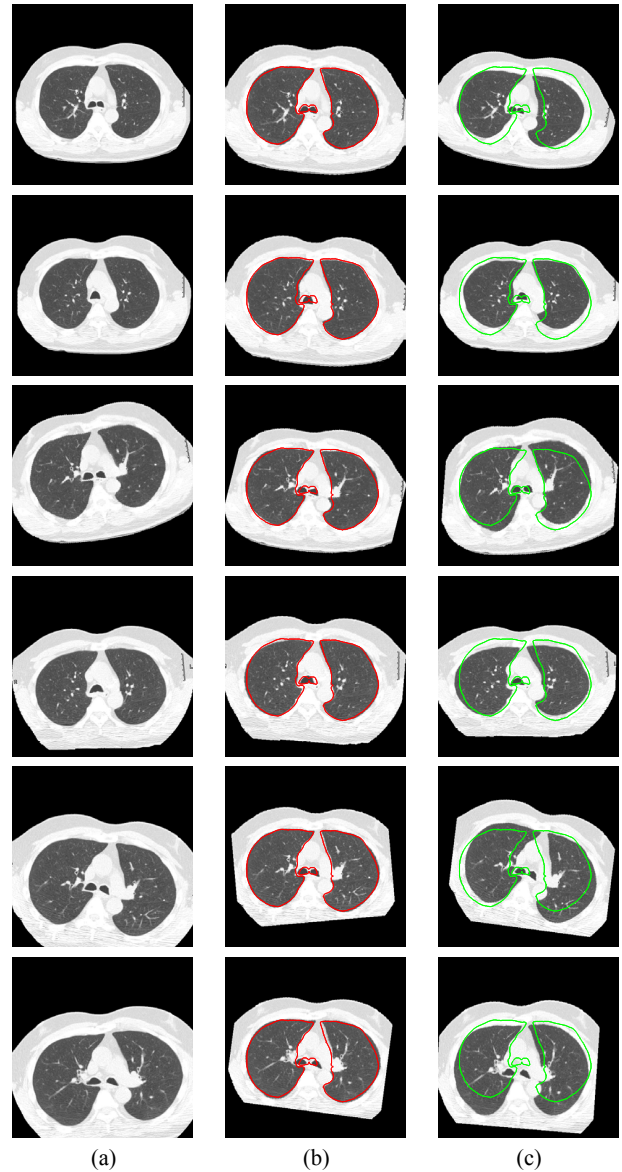
#### 5. REFERENCES

- [1] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, issue 1, pp. 1–36, 1998.

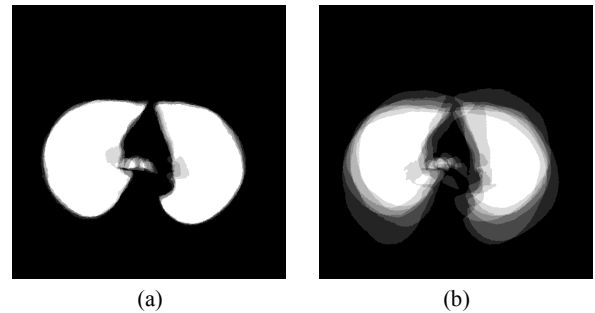


**Fig. 5.** Gibbs energy, MI, and NMI values at the successive steps of the gradient search.

- [2] P. Hellier, "Inter-subject non-rigid registration: survey and methods," In *Handbook of medical image analysis: segmentation and registration methods*, (J. Suri, D. Wilson, S. Laxminarayan, Eds.), Vol. 3, , Kluwer, 2004.
- [3] X. Huang, Y. Sun, D. Metaxas, F. Sauer, and C. Xu, "Hybrid Image Registration based on Configurational Matching of Scale-Invariant Salient Region Features," *IEEE Workshop on Image and Video Registration (IVR04) in conjunction with CVPR04*, Washington D.C., July, 2004.
- [4] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.
- [5] Pope and J. Theiler, "Automated Image Registration (AIR) of MTI Imagery," *Proc. SPIE 5093*, vol. 27, pp. 294–300, 2003.
- [6] J. Inglada and A. Giros, "On the Possibility of Automatic Multi-sensor Image Registration," *IEEE Trans. Geoscience and Remote Sensing*, vol. 42, no. 10, October 2004.
- [7] P. Viola, "Alignment by Maximization of Mutual Information," *Ph.D. dissertation, MIT*, Cambridge, MA, 1995.
- [8] P. Thevenaz and M. Unser, "Alignment An efficient mutual information optimizer for multiresolution image registration," *Proc. IEEE Int. Conf. on Image Processing ICIP '98*, Chicago, IL, 1998, pp. 833–837.
- [9] P. Thevenaz and M. Unser, "A pyramid approach to sub-pixel image fusion based on mutual information," *Proc. IEEE Int. Conf. on Image Processing ICIP '96*, Lausanne, Switzerland, 1996, pp. 265–268.
- [10] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–86, 1999.
- [11] J. Pluim, J. Maintz, and M. Viergever, "Mutual-Information-Based Registration of Medical Images: A Survey," *IEEE Trans. Medical Imaging*, vol. 22, no. 8, August, 2003.
- [12] A. El-Baz, "Novel Stochastic Models for Medical Image Analysis," *Ph.D. dissertation, University of Louisville*, Louisville, KY, 2006.
- [13] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. of Computer Vision*, vol. 60, pp. 91–110, 2004.



**Fig. 6.** Original lungs (a) aligned with our (b) and the MI-based (c) approaches.



**Fig. 7.** Overlap between the object masks aligned with our (a; 96.8%) and the MI-based approaches (b; 54.2%).